MOLECULAR AND CELLULAR BIOLOGY, Aug. 1997, p. 4820–4829 Vol. 17, No. 8
0270-7306/97/$04.00+0


# Two Evolutionarily Conserved Repression Domains in the *Drosophila Krüppel* Protein Differ in Activator Specificity

WENDY HANNA-ROSE,[1]† JONATHAN D. LICHT,[2] AND ULLA HANSEN[1]*

*Division of Molecular Genetics, Dana-Farber Cancer Institute and Department of Microbiology and Molecular Genetics, Harvard Medical School, Boston, Massachusetts 02115,[1] and Brookdale Center for Molecular Biology and Department of Medicine, Mount Sinai School of Medicine, New York, New York 10029[2]*

Received 13 February 1997/Returned for modification 9 April 1997/Accepted 29 April 1997

**To identify biologically functional regions in the product of the *Drosophila melanogaster* gene *Krüppel*, we cloned the *Krüppel* homolog from *Drosophila virilis*. Both the previously identified amino (N)-terminal repression region and the DNA-binding region of the *D. virilis Krüppel* protein are greater than 96% identical to those of the *D. melanogaster Krüppel* protein, demonstrating a selective pressure to maintain the integrity of each region during 60 million to 80 million years of evolution. An additional region in the carboxyl (C) terminus of *Krüppel* that was most highly conserved was examined further. A 42-amino-acid stretch within the conserved C-terminal region also encoded a transferable repression domain. The short, C-terminal repression region is a composite of three subregions of distinct amino acid composition, each containing a high proportion of either basic, proline, or acidic residues. Mutagenesis experiments demonstrated, unexpectedly, that the acidic residues contribute to repression function. Both the N-terminal and C-terminal repression regions were tested for the ability to affect transcription mediated by a variety of activator proteins. The N-terminal repression region was able to inhibit transcription in the presence of multiple activators. However, the C-terminal repression region inhibited transcription by only a subset of the activator proteins. The different activator specificities of the two regions suggest that they repress transcription by different mechanisms and may play distinct biological roles during *Drosophila* development.**

The spatial and temporal control of eukaryotic gene transcription depends on the interplay of numerous positive and negative *trans*-acting factors. Transcriptional activator proteins and their modes of action have been extensively characterized (reviewed in reference 49). In contrast, although regulation by repression is critically important for maintaining appropriate levels of gene expression, only recently have the molecular mechanisms of repressors come under significant scrutiny (reviewed in references 6, 19, 22, and 36).

Like transcriptional activators (2), many transcriptional repressors are modular factors (27), in that the transcriptional effector domain can be separated from the DNA-binding domain. The activation regions of many transcriptional activators have been identified and categorized based on primary amino acid sequence homology (49). Although less well defined to date, transferable repression domains have been identified in several transcriptional repressors. Like activation domains, some repression domains fall into distinct classes based on primary amino acid sequence composition, presumably representing repressors that function through different sets of protein-protein interactions.

Activation domains fall into classes including glutamine-rich, proline-rich, and acidic domains. For many transcription factors in each class, maximal activation appears to be the result of multiple interactions between the activator and targets within the transcriptional machinery (reviewed in reference 49). Direct interactions with components of the basal transcriptional machinery as well as indirect, coactivator-mediated interactions have been implicated in transcriptional activation. Multiple classes of activators appear to interact with specific basal transcription factors. Both VP16 (45), a representative member of the acidic activator class, and Sp1 (8), a representative member of the glutamine-rich class of activators, interact specifically with the human TATA-binding protein. In contrast, different classes of activators appear to contact distinct TATA-binding protein-associated factors (TAFs) and unique coactivator proteins to activate transcription. For example, the acidic activators VP16, p53, and Epstein-Barr virus nuclear protein 2 interact with *Drosophila* $TAF_{II}40$ (15, 48, 50), whereas the glutamine-rich activation domains of Sp1 and the glutamine-rich activation domain, Q2, of the cyclic AMP-regulated transcription factor CREB interact with *Drosophila* $TAF_{II}110$ (10, 13, 20).

Classifications based entirely on amino acid sequence similarity are preliminary. In fact, the amino acids used to characterize an activation domain are not necessarily the residues that contribute to activation function in the region. The acidic residues in the VP16 activation domain (35) and glutamine residues in the Sp1 activation domain (13) can be mutagenized without affecting activation function. In contrast, hydrophobic residues appear to play a larger role in activation activity in the case of both proteins (13, 35).

Repression domains can also be classified according to primary amino acid sequence to include charged, alanine-rich, and proline-rich domains (reviewed in reference 19). As is sometimes true for classes of activation regions, separate classes of repression regions may utilize distinct mechanisms to regulate transcription. Proposed mechanisms for negative regulation of transcription include both passive mechanisms, involving steric hindrance, and active mechanisms, involving protein-protein interactions between the repressor and its target, such as direct repression and quenching (19, 26). Direct repression involves inhibition of basal transcription, whereas quenching involves interference with the function of a specific

* Corresponding author. Mailing address: Dana-Farber Cancer Institute, Room B411, 44 Binney St., Boston, MA 02115. Phone: (617) 375-8211. Fax: (617) 375-8221. E-mail: ulla_hansen@dfci.harvard.edu.

† Present address: Department of Molecular, Cellular and Developmental Biology, University of Colorado, Boulder, CO 80309-0347.

activator protein. One indication of the potential mechanism of a particular repressor is its ability to inhibit all or only a subset of promoters or to inhibit transcription mediated by specific activators and not others. A repressor that functions in all, or most, contexts is more likely to inhibit basal transcription. In contrast, a repressor that can inhibit transcription only in particular contexts is more likely to target a specific activator(s) or proteins that specifically mediate the action of an activator.

The first example of a transferable repression domain was found in the amino (N) terminus of the product of the *Drosophila melanogaster* gap gene *Krüppel* (*Kr*) (27). Kr is a sequence-specific, DNA-binding protein (34, 34a, 44, 52). When bound upstream of an appropriate RNA polymerase II promoter, Kr represses transcription from that promoter in both mammalian and *Drosophila* tissue culture systems (27, 40, 57). The minimal N-terminal repression domain of Kr, the smallest and best defined of repression domains, consists of a probable α helix, encompassing residues 62 to 92, and contains a glutamine at residue 86 that is critical for repression function (28). Previously, the extents of transcriptional repression by Kr were compared in the presence of two different activator proteins. Whereas full-length Kr repressed transcription in the presence of Gal4-Q, a fusion of the Gal4 DNA-binding domain to the second glutamine-rich activation domain of Sp1, it was ineffective in the presence of the wild-type acidic activator Gal4. These data suggested that Kr mediates repression through a quenching mechanism (29). In addition, repression by Kr has been correlated with an interaction with the basal transcription factor TFIIE (39).

In the current study, by examining the conservation of *Kr* during *Drosophila* speciation, we show that the three most highly conserved portions of the gene include the region encoding the zinc finger DNA-binding motif, the region in the N terminus of the gene encoding the previously identified N-terminal repression region, and a region in the carboxyl (C) terminus of the protein. The C-terminal 101 amino acids had previously been shown to contain a transferable repression region in *Drosophila* cells; we now demonstrate that a 42-amino-acid region within the C terminus encodes a transferable repression region of equivalent activity, which is also active in mammalian cells. The smaller C-terminal repression region has a modular primary amino acid structure consisting of a basic subregion, a proline-rich subregion, and an acidic subregion. We have used a mutagenesis strategy to analyze the contributions of each subregion to repression function and demonstrated that the acidic residues contribute to transcriptional repression.

To investigate the mechanistic bases for repression by Kr, we have extended the original activator specificity studies by testing the ability of each of the two Kr repression regions to attenuate transcription mediated by various activator proteins. The N-terminal repression region inhibited transcription mediated by activators against which the C-terminal repression region was ineffective. These data suggest that the two repression regions of Kr repress transcription via different mechanisms and have distinct biological functions during *Drosophila* development.

## MATERIALS AND METHODS

**Plasmid constructions.** During the course of this work, we found that a single guanine residue in codon 410 was omitted from the published sequence of the *D. melanogaster Kr* cDNA, pcK2b (37) (Fig. 1). Inclusion of this residue results in the corrected open reading frame of Kr being 502 amino acids; this new numbering is reflected in the constructions described below. All Kr expression plasmids express the *D. melanogaster* protein. Kr-NZC (CMV-Kr), Kr-NZΔ (CMV-KrΔC), Kr-NZΔ•86QK (CMV-KrΔC•86QK), CMV-LacI/Z (12, 28), and the reporter genes pSV*lacO*CAT (3), pSVCAT (3), Kr4-tkCAT (27), G5BCAT (5), and Kr4G5BCAT (29) have been described previously. The activator expression plasmids Gal4-Q (29), Gal4 (pSG4) (23), and Gal4-AH (pSG50.1) (14) were also described previously. Gal4-HOX contains the first 68 amino acids of the mouse *Hox1.3* gene (11, 56) fused to Gal4 residues 1 to 147. The reporter plasmid Lac2-tkCAT was constructed by inserting two copies of a symmetrical *lac* operator sequence (12) with high affinity for LacI into the *Sal*I site of pBLCAT2 (30), upstream of the herpes simplex virus (HSV) thymidine kinase (*tk*) promoter. Kr-ΔZC, which lacks sequences encoding amino acids 31 to 214 of Kr, was constructed by digestion of CMV-Kr (28) with *Xba*I and *Bam*HI and religation of the plasmid in the presence of an adapter linker with the sequence

```
5′ GATCACGGT      3′
3′      TGCCAGATC 5′
```

to maintain the correct reading frame. Lac/Kr-C(345-502) was constructed by digestion of pcK2b with *Ban*I and *Eco*RI and ligation of this fragment to a 1.4-kb *Bgl*II-*Pvu*I fragment, encoding the *lac* repressor linked to 23 N-terminal amino acids of β-galactosidase, using the synthetic adapter

```
5′   CTCGAGCGGCG      3′
3′ TAGAGCTCGCCGCCGTG 5′
```

This fusion gene was cloned into pBluescript II SK+ (Stratagene), excised by digestion with *Sac*I and *Eco*RI, and cloned into construct CMV-LacI digested with *Sac*I and *Eco*RI. Lac/Kr-C(402-502) was constructed by digestion of Lac/Kr 26-466(502) (27) with *Xho*I and *Eco*RI to isolate a backbone encoding the LacI/Z fusion protein. Next, a *Sal*I-*Eco*RI fragment from Lac/Kr 26-466(502) encoding amino acids 402 to 502 of Kr was isolated and ligated to the backbone. Lac/Kr-C(433-474) was created by amplifying *Kr* sequences with a stop codon, using the primers 5′ CTCGAGCTCGAGGCACGGCGCAAGGCACAG 3′ and 5′ GGA TCCGGATCCCTAAGAACGAGGAGAGTGCAT 3′, digesting the resulting fragment with *Xho*I and *Bam*HI, and ligating it to the backbone from CMV LacI/Z VP16 (28) digested with *Xho*I and *Bam*HI. All amplified and subcloned fragments were sequenced to ensure that no mutations had been generated and that an in frame fusion with the *lacI/Z* gene fusion was created.

**Mutagenesis.** The C-terminal multiple point mutations in Kr were created by oligonucleotide site-directed mutagenesis using the Sculptor in vitro mutagenesis system (Amersham Life Sciences). Utilizing the VCS-M13 helper phage (Stratagene), we produced single-stranded DNA from a pBluescript II SK+ vector into which the *Xho*I-*Bam*HI fragment from CMV Lac/Kr-C(433-474) had been ligated. The mutant Lac/Kr-C(433-474•K/R/D/E-A) was created by producing single-stranded DNA from the Lac/Kr-C(433-474•K/R-A) mutant and performing a second round of mutagenesis. The mutated constructs were sequenced to ensure the absence of undesired, additional mutations.

**Cell culture and transfections.** CV-1 (African green monkey kidney) cells and U2OS (human osteosarcoma) cells were grown in a 5% $CO_2$ environment in Dulbecco's modified Eagle's medium supplemented with 10% calf serum or 10% fetal bovine serum, respectively. Cells were plated at a density of $7 \times 10^5$ to $1 \times 10^6$ cells/100-mm-diameter dish 24 h prior to transfection, and transfections performed by the calcium phosphate precipitation method were done as described previously (25, 27, 28). Reporter and effector plasmids were transfected in amounts indicated in the figure legends along with either pMTGH or tkGH, both of which are human growth hormone (hGH) expression plasmids (43), as an internal control of transfection efficiency, and either pBluescript II SK+ (Stratagene) or pUC18 as filler DNA to maintain 10 to 15 μg of DNA in each transfection. In U2OS cells, the optimal conditions for obtaining DNA-binding site-dependent repression occurred at much lower ratios of effector plasmid to reporter plasmid than in CV-1 cells. Transfections with the pSV*lacO*CAT and pSVCAT reporters contained 1.0 μg of reporter plasmid and 0.5 μg of effector plasmid.

To transfect cells by using LipofectAMINE, CV-1 cells (approximately $2 \times 10^5$ to $3 \times 10^5$) were plated in 60-mm-diameter tissue culture dishes such that they were 50% confluent after 24 h. Effector and reporter plasmids were transfected in the amounts indicated in the figure legends along with 0.5 μg of tkGH as an internal control of transfection efficiency and pBluescript II SK+ (Stratagene) as filler DNA to maintain 8 μg of DNA in each transfection. Plasmid DNAs mixed in 300 μl of cell culture medium (without serum) were combined with 20 μl of LipofectAMINE reagent (GibcoBRL) mixed in 300 μl of cell culture medium (without serum). Complexes were allowed to form for 30 to 40 min at room temperature. Cells were washed twice and overlaid with 2.4 ml of culture medium (without serum). The LipofectAMINE-DNA mixture was added to the cells and incubated at 37°C in a $CO_2$ incubator for 5 h. Medium was replaced with 5 ml of cell culture medium containing 10% calf serum after 5 h and again after 24 h.

At 48 h posttransfection by either method, cell culture media were harvested for hGH radioimmunoassay (Allegro, San Capistrano, Calif.) as instructed by the manufacturer. In addition, cytoplasmic extracts were prepared as described previously (16, 25, 27), and equal amounts of protein, as determined by the Bradford assay (Bio-Rad), were assayed for chloramphenicol acetyltransferase (CAT) activity. The percent conversion of chloramphenicol to acetylated chloramphenicol was determined by analysis of chromatograms with a PhosphorImager and ImageQuant software (Molecular Dynamics, Sunnyvale, Calif.). Corrected CAT activity is defined as the percent conversion of chloramphenicol to acetylated

FIG. 1. Nucleotide sequence comparison of the *D. virilis* and *D. melanogaster Kr* genes. The numbered sequence of the *D. melanogaster* cDNA clone pcK2b (37), beginning with the AUG initiation codon, is shown on the second lines, and the sequence of the *D. virilis* genomic clone is shown above. Aligned nucleotides are indicated with capital letters and unaligned nucleotides are in lowercase. Dots indicate gaps, and dashes indicate identities. The carat flanked by *D. melanogaster* nucleotide positions 37 and 410 indicates the position of the intron (not shown here) in the *D. melanogaster* gene. Boldface underlined nucleotides in the *D. melanogaster* sequence indicate amendments to the published sequence of pcK2b: position 576 was amended to C from A, position 783 was amended to C from T, and the G at position 1602 is an addition to the sequence of pcK2b, leading to a shift in the reading frame of the encoded protein (see Fig. 2). Due to the *Sal*I site in the 5′ coding region of the *D. virilis* gene, the *Sal*I fragment subclone that was sequenced does not include the extreme N terminus of the *D. virilis* coding region.

chloramphenicol divided by the level of hGH (nanograms per milliliter) obtained in the same transfection experiment. Relative CAT activities were determined by normalizing the corrected CAT activity for each effector to the corrected CAT activity in the presence of the control LacI/Z protein (set at 100% activity).

**Cloning of the *Drosophila virilis Kr* gene.** To clone the *D. virilis Kr* gene, a *D. virilis* library of genomic DNA partially digested with *Mbo*I in an EMBL3 vector (gift of Ronald K. Blackman) was screened with a 561-bp probe obtained by digesting the *D. melanogaster Kr* cDNA clone, pcK2b, with *Xba*I and *Bam*HI and labeling by random priming. Sixty thousand plaques were plated, lifted onto duplicate filters, prehybridized overnight, and then hybridized in 43% formamide–5× SSPE (0.9 M NaCl, 50 mM $NaH_2PO_4$–5 mM EDTA [pH 7.7])–5× Denhardt's solution (0.1% [wt/vol] Ficoll 400, 0.1% [wt/vol] polyvinylpyrrolidone, 0.1% [wt/vol] bovine serum albumin)–1% sodium dodecyl sulfate–100 μg of denatured salmon sperm DNA per ml with 106 cpm of probe per ml of hybridization solution at 37°C overnight. The filters were washed in 2× SSPE–1% sodium dodecyl sulfate for 1 h at 55°C and then for 1 h at 60°C. Eleven positive plaques were picked for secondary screening. DNA was prepared from seven positive clones, and Southern analysis (32) using the same probe and hybridization conditions as above was performed. Each blot was stripped and rehybridized with a probe coding for the zinc finger region of the Kr protein, obtained by digesting pcK2b with *Eco*RI and *Bam*HI. Only one phage, I3.3, contained a fragment that hybridized to both probes. The 4.4-kb *Sal*I fragment from I3.3 that hybridized to both probes was subcloned into pBluescript II SK+ (Stratagene) and sequenced on both strands.

**Nucleotide sequence accession number.** The GenBank accession number for the sequence of the *D. virilis Kr* clone is U49856.

## RESULTS

**Three specific regions of Kr were highly conserved during *Drosophila* speciation.** To investigate the biologically relevant regions of the *Drosophila* Kr protein, we cloned the *D. melanogaster Kr* homolog from *D. virilis*. These two species of *Drosophila* are separated by 60 million to 80 million years of evolution, a sufficient time for complete divergence of unconstrained sequences (1, 24, 46). Conserved sequences, therefore, likely represent essential and functional regions of the protein homologs (1, 24, 46). To obtain the *D. virilis* gene, a *D. virilis* genomic library was screened at low stringency, using a probe encoding the N terminus but not the zinc finger region of Kr, thus preventing isolation of the large family of zinc finger genes in the *Drosophila* genome. Of the seven positive clones identified, only one hybridized with a fragment from the *Kr* cDNA that encodes the zinc finger and the C terminus of the *D. melanogaster* Kr protein. A 4.4-kb DNA fragment that hybridized to both probes was subcloned and sequenced on both strands. Comparisons of the nucleotide and the amino acid sequences to those of the *D. melanogaster* gene and protein are shown in Fig. 1 and 2A, respectively. The presented *D. melanogaster* nucleotide and amino acid sequences differ slightly from the previously published sequences, as discussed below.

The degree of homology between the two genes at the nucleotide level is high (66.9% identical) throughout the coding regions. Half of the divergence is due to changes at the third position of the codons. The homology drops off sharply beyond

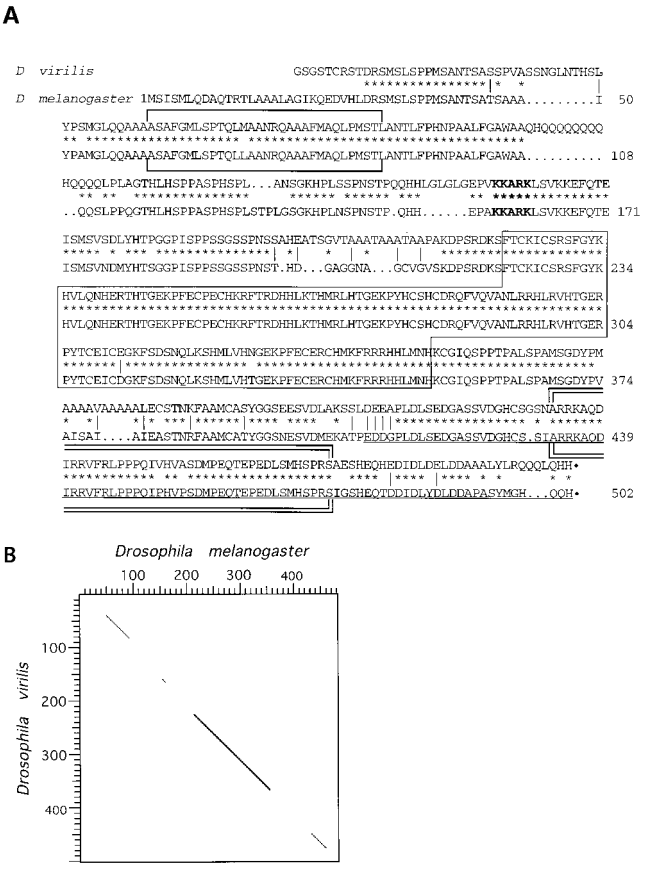


FIG. 2. Amino acid sequence comparison of the *D. virilis* and *D. melanogaster* Kr proteins. (A) The predicted amino acid sequence of the *D. melanogaster* gene, shown on the bottom line, is compared to that of the *D. virilis* gene above. *, identical amino acids; |, conservative amino acid substitution; •, stop codon; and ., gap in the alignment of the two sequences. Brackets delineate the minimal N-terminal repression domain, double brackets delineate the minimal C-terminal repression domain, and the box outlines the zinc finger region. The nuclear localization sequence is in boldface. The underlined portion of the *D. melanogaster* protein differs from the previously published sequence due to a correction of the reading frame (see Fig. 1). (B) Dot matrix homology analysis (DNA Strider 1.2) of the Kr proteins from *D. melanogaster* and *D. virilis*. *D. melanogaster* amino acid sequence numbers are on the horizontal axis, and *D. virilis* amino acid sequence numbers (using the initial glycine in Fig. 2A as position 1) are on the vertical axis. The window used for the search was 23 amino acids, with a maximum of only two mismatches allowed.

the 3′ end of the coding region (only 27.5% identity). The aligned *D. virilis* coding sequence is 120 nucleotides longer than that of *D. melanogaster* due to a series of six short insertions relative to the *D. melanogaster* gene. These occur at regions of cryptic simplicity (i.e., short, partly irregular, direct repeats) that are predicted to be prone to a slippage mechanism of mutagenesis (47, 51).

The high degree of homology at the nucleotide level is reflected in the 81% identity of the two predicted amino acid sequences (Fig. 2A). Note that three regions of the protein are nearly completely conserved, demonstrating selective pressure to maintain the integrity of each region during evolution (Fig. 2B). A fourth, very small segment of conservation (Fig. 2B and highlighted in Fig. 2A) corresponds to the position of the nuclear localization sequence of the Kr protein (21). The largest region of conservation includes 140 amino acids encompassing the zinc finger DNA-binding domain, which is 98.6% identical between the two species. Only two amino acid substitutions, one a conservative substitution of a glutamic acid for an aspartic acid, interrupt the identity of this region. A second region of 95.5% identity lies in a 44-amino-acid, N-terminal region of the protein, encompassing the previously identified N-terminal repression domain (Fig. 2). Only one amino acid substitution alters the minimal N-terminal repression region of *D. virilis* relative to *D. melanogaster*. This finding demonstrates the importance of this region to the biology of *Drosophila* and suggests that the definition of the repression domain in mammalian cells relates to the function of the protein during *Drosophila* embryogenesis. The third region of near perfect conservation lies in the C terminus of the protein. A 62-amino-acid region, identified in the dot homology plot analysis, is 93% identical between the two species (Fig. 2A and B). This region is contained within the 101-amino-acid repression region identified previously in *Drosophila* tissue culture systems (40, 57) and is further defined below.

The sequence of the predicted *D. virilis* Kr protein was distinctly similar to that of the previously published *D. melanogaster* protein through amino acid 403, at which point the homology abruptly ended. However, when the predicted C-terminal sequence of the *D. virilis* Kr protein was compared to a different reading frame in the published *D. melanogaster Kr* cDNA sequence, a striking match could be made. To determine if this apparent reading frame shift in the *D. virilis* gene relative to the *D. melanogaster* gene was real or the result of an error in either sequence, genomic DNA from both species was sequenced in this region. Sequencing of a separate source of *D. virilis* genomic DNA confirmed the sequence shown in Fig. 1 (data not shown). However, when analyzed under appropriate gel electrophoretic conditions, the *D. melanogaster* genomic DNA contained an extra base pair in codon 410 relative to the published sequence of the *D. melanogaster* cDNA, pcK2b (data not shown). The same region of pcK2b was subsequently sequenced and found to agree with our genomic sequence rather than the published sequence. Similar results have been obtained in other laboratories (21). The additional guanine nucleotide at position 1602 (Fig. 1) results in the alteration of the published reading frame in the C-terminal portion of the *D. melanogaster* protein and brings it into register with the *D. virilis* clone. The *D. melanogaster* Kr protein accordingly contains 502 amino acids, with a predicted molecular weight of 54,700 (Fig. 1), rather than the previously stated 466 amino acids. The correct stop codon is at nucleotide position 1879. This also alters the identity of the C-terminal repression region as assayed in *Drosophila* Schneider cells (40). Although the repression region was stated to be within the C-terminal 64 amino acids of Kr (40), after correction of the reading frame, this activity actually lies within the last 101 amino acids of the protein.

**The highly conserved region in the C-terminal portion of the Kr protein contains a transferable repression activity.** We previously showed that an N-terminal region of the Kr protein could confer repression activity on the LacI protein when it was fused to that DNA-binding protein (27). However, subsequent experiments indicated that deletion of the N terminus from the Kr protein resulted in a derivative that repressed transcription of a promoter containing Kr DNA-binding sites to approximately the same extent as the full-length protein (8a). This finding suggested the possibility of redundant repression regions in the Kr protein that are active in mammalian cells, one in the N terminus and the other in the C terminus. To identify the C-terminal repression activity, portions of the C terminus of Kr were fused to the *lacI* DNA-binding protein and assayed in U2OS and CV-1 cells by cotransfection with the reporter construct Lac2-tkCAT, containing two *lac*

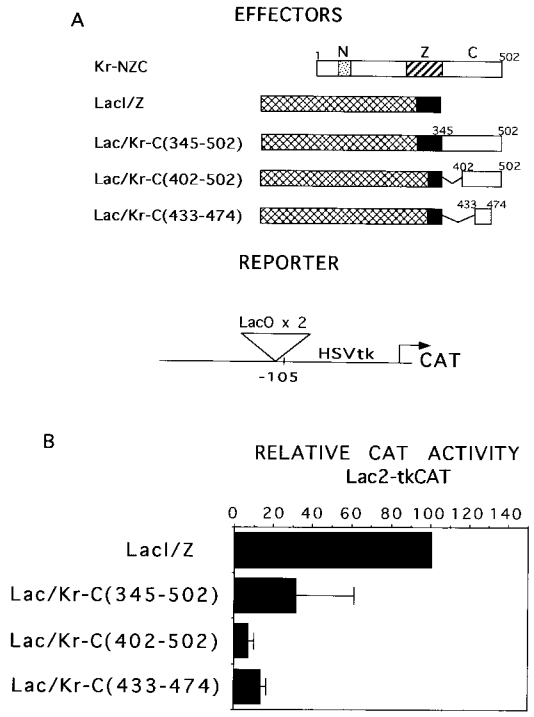


FIG. 3. The C-terminal repression domain of Kr maps to a small, highly conserved region. (A) Diagrams of the control and Lac/Kr effector constructs and of the reporter construct. All proteins were expressed in U2OS cells under the control of the simian cytomegalovirus promoter. The hatched region indicates the *lacI* gene and the solid region indicates the portion of *lacZ* in the constructs. The C-terminal regions of Kr in the fusions are indicated with open boxes and marked C; the diagonally striped region of full-length Kr indicates the zinc finger DNA-binding region and is marked Z; the shaded region indicates position of the previously identified N-terminal repression region and is marked N. (B) Repression of Lac2-tkCAT gene expression by Lac/Kr-C fusion proteins. Lac2-tkCAT (3 μg) was cotransfected into U2OS cells with 0.5 μg of each of the Lac/Kr-C fusion expression plasmids or a control expression plasmid (CMV-LacI/Z) and 2 μg of internal control tkGH plasmid, using a calcium phosphate transfection protocol. Relative CAT activity was determined by normalizing levels of corrected CAT activity to the percentage of full activity in the presence of the control LacI/Z plasmid. The indicated activity was determined by averaging activities from four independent experiments performed in duplicate; standard deviations are indicated.

operators located upstream of the HSV *tk* promoter (Fig. 3A). Qualitatively similar data were obtained for both cell lines, although the degree of repression was three- to fivefold higher in U2OS cells. Lac/Kr-C(345-502) and Lac/Kr-C(402-502) each mediated repression of CAT gene expression directed by Lac2-tkCAT (Fig. 3B). These data indicated that the C-terminal 101 amino acids of the Kr protein could function as a repression domain in mammalian cells, as was previously observed in *Drosophila* Schneider cells. Given the strong homology with the *D. virilis* protein between residues 414 and 474 (Fig. 2B), we predicted that the C-terminal repression domain would lie in this conserved region. Indeed, a fusion protein between amino acids 433 and 474 of Kr and the *lacI* DNA-binding protein, Lac/Kr-C(433-474), repressed expression from Lac2-tkCAT as well as the entire C terminus fused to LacI, Lac/Kr(345-502) (Fig. 3B). A fusion of the *lacI* DNA-binding protein to amino acids 400 to 437 of Kr was also tested and did not repress transcription from Lac2-tkCAT (data not shown). However, this protein was not expressed in U2OS or CV-1 cells (see below); thus, its repression activity could not be evaluated.

To ensure that all of the Lac/Kr fusion proteins were expressed, localized to the nucleus, and able to specifically bind DNA, each construct was cotransfected with the reporter pSV*lacO*CAT (12). pSV*lacO*CAT contains a *lac* operator directly over the start site of the simian virus 40 (SV40) early promoter. Gene expression from this construct is inhibited by the mere expression of a protein containing the *lacI* DNA-binding domain, due to steric hindrance between the binding of the LacI fusion protein and the RNA polymerase II transcription complex at the transcriptional start site. This type of repression is in stark contrast to repression of the Lac2-tkCAT reporter, discussed above, in which the *lacI* recognition sites are positioned over 100 bp upstream of the start site. Transcriptional repression from such upstream sites requires expression of a DNA-binding protein fused to a transferable repression region. To verify that effects of LacI fusion proteins on pSV*lacO*CAT expression were specific, repression of pSV*lacO*CAT was compared to repression from pSVCAT, a reporter containing the SV40 early promoter but no *lac* operator. Both constructs were analyzed in the presence of the positive control LacI/Z or various Lac/Kr fusions. As expected, full-length Kr did not repress either promoter, because it cannot bind to a *lac* operator. In contrast, LacI/Z and each Lac/Kr fusion, with one exception, specifically repressed pSV*lacO*CAT expression, with quantitative effects ranging to between 0.28 and 3.5% of the control level of expression (Table 1). This analysis demonstrated that these fusion proteins were expressed and active in the nuclei of the transfected cells. Only Lac/Kr-C(400-437) did not repress pSV*lacO*CAT expression, indicating that this fusion is either not expressed in the nucleus of the cell or not capable of binding to DNA.

**Acidic residues in the C terminus of the Kr protein contribute to repression activity.** The 42-residue, C-terminal repression domain (residues 433 to 474) shows no sequence or structural homology to the N-terminal repression domain. However, the C-terminal region does exhibit a striking modularity in its primary amino acid sequence (Fig. 4A). The N-terminal portion of this short region has a high proportion of basic residues, the central portion has a high proportion of proline residues, and the C-terminal portion has a high proportion of acidic residues. To examine the involvement of these subregions in repression function, each was individually targeted by oligonucleotide site-directed mutagenesis. Three multiple alanine substitution mutants were constructed: Lac/Kr-C(433-474•K/R-A), which targets each basic residue in the N-terminal subregion; Lac/Kr-C(433-474•P-A), which targets

TABLE 1. Effects of different Lac/Kr fusion proteins on expression from the cotransfected pSV*lacO*CAT

| LacI/Z, Kr, or Lac/Kr fusion protein expressed upon transfection | CAT expression from pSV*lacO*CAT (% of control [mean ± SD]) |
|---|---|
| LacI/Z | 0.28[a] |
| Kr | 100 ± 4.9 |
| Lac/Kr-C(345-502) | 2.56 ± 0.14 |
| Lac/Kr-C(402-502) | 3.49 ± 0.08 |
| Lac/Kr-C(400-437) | 124 ± 4.5 |
| Lac/Kr-C(433-474) | 0.95 ± 0.02 |
| Lac/Kr-C(433-474•K/R-A) | 1.33 ± 0.02 |
| Lac/Kr-C(433-474•P-A) | 2.18 ± 0.04 |
| Lac/Kr-C(433-474•D/E-A) | 1.99 ± 0.04 |
| Lac/Kr-C(433-474•K/R/D/E-A) | 0.62 ± 0.02 |

[a] The assay was performed only once, as a positive control; thus, no standard deviation is given.

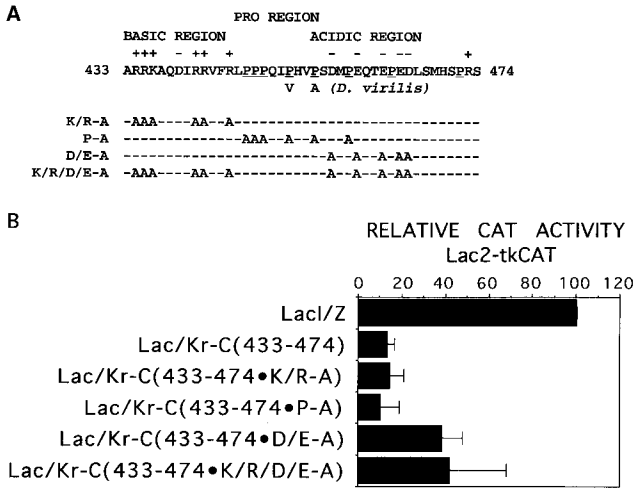


FIG. 4. Multiple point mutations in the C-terminal repression region of the Kr protein. (A) Amino acid sequence of the minimal C-terminal repression region. The three subregions are indicated. Basic residues are marked above by +. Proline residues are underlined, and acidic residues are marked above by −. Residues that are not conserved in the *D. virilis* protein are indicated with the substitution below the sequence. The sequence of each of the multiple alanine substitution mutants is given. (B) Repression of Lac2-tkCAT gene expression by Lac/Kr-C(433-474) mutants. The indicated effector constructs were transfected into U2OS cells as described in the legend to Fig. 3B. Relative CAT activity was determined by normalizing levels of corrected CAT activity to the percentage of full activity (determined as percent conversion of chloramphenicol) in the presence of the control LacI/Z plasmid. The indicated activity was determined by averaging activities from four independent experiments performed in duplicate; standard deviations are indicated.

each proline residue in the central subregion; and Lac/Kr-C(433-474•D/E-A), which targets each acidic residue in the C-terminal subregion (Fig. 4A). Each protein was assayed for repression activity by cotransfection of its expression construct and the Lac2-tkCAT reporter in U2OS cells. In addition, each expression construct was assayed for appropriate expression of active protein by cotransfection with the reporters pSV*lacO*CAT and pSVCAT. Lac/Kr-C(433-474) and the mutant fusion proteins specifically repressed pSV*lacO*CAT gene expression to between 0.62 and 2.2% of control levels (Table 1), demonstrating that each protein was present and competent to bind DNA in the nuclei of transfected cells.

Mutation of either the basic subregion or the proline-rich subregion region did not diminish the ability of the C-terminal repression region to inhibit expression from Lac2-tkCAT (Fig. 4B). However, mutation of the acidic subregion resulted in a protein that repressed Lac2-tkCAT expression 2.8-fold less than Lac/Kr-C(433-474) (Fig. 4B). Because Lac/Kr-C(433-474•D/E-A) was expressed at somewhat lower levels than Lac/Kr-C(433-474), as assayed by the pSV*lacO*CAT cotransfection experiments (Table 1), we constructed a fourth mutant both to confirm the role of the acidic residues in transcriptional repression and to investigate the possible redundancy of the acidic and basic residues in inhibiting transcription. All of the basic residues in the first subregion and the acidic residues in the third subregion were simultaneously changed to alanine in the fourth mutant, Lac/Kr-C(433-474•K/R/D/E-A) (Fig. 4A). Although at least as active as Lac/Kr-C(433-474•K/R-A) in inhibiting expression of pSV*lacO*CAT (Table 1), Lac/Kr-C(433-474•K/R/D/E-A) repressed Lac2-tkCAT expression 3.0-fold less than Lac/Kr-C(433-474•K/R-A) (Fig. 4B). The diminished repression activity of Lac/Kr-C(433-474•D/E-A) compared to the wild-type repression activity of Lac/Kr-C(433-474) and the diminished repression activity of Lac/Kr-C(433-474•K/R/D/E-A) compared to the wild-type repression activity of Lac/Kr-C(433-474•K/R-A) indicate that the acidic residues do play a role in inhibiting gene expression from upstream sites. Therefore, surprisingly, mutation of the basic or proline-rich subregions individually did not affect the transferable repression activity of the C-terminal repression region of *Kr*, whereas mutation of the acidic residues did.

**Gal4-Q- or Gal4-AH-mediated transcription is repressed by the N-terminal and the C-terminal repression regions of Kr.** To probe whether the N-terminal and C-terminal repression regions of Kr are functionally redundant, or whether they might instead play distinct biological roles, we investigated their abilities to inhibit transcription mediated by various transcriptional activators. The activator specificity assay is outlined in Fig. 5. In brief, an expression vector for Kr or a Kr deletion derivative, or the same expression vector for a nonbinding control protein, LacI/Z, was cotransfected along with an expression vector for an activator protein and a synthetic CAT reporter construct. The promoter region driving CAT gene expression contains multiple binding sites for both Gal4 and Kr upstream of a TATA sequence (Fig. 5A). Various activator regions, all containing the DNA-binding domain of Gal4 (Fig. 5B), were used to promote expression of CAT. The ability of each Kr derivative to inhibit CAT gene expression mediated by each specific activator was then determined. Because the DNA-binding domain of each activator is a constant, we can specifically probe the ability of each repression region to inhibit different types of activation pathways. We tested three repressors: Kr-NZC, Kr-ΔZC, and Kr-NZΔ (Fig. 5B). Kr-NZC represents the full-length Kr protein and contains both the N-terminal and C-terminal repression regions. Kr-NZΔ lacks the C-terminal 155 amino acids of Kr, retaining only the N-terminal repression region. Conversely, Kr-ΔZC lacks the 184 N-terminal amino acids of Kr, retaining only the C-terminal repression region. As a negative control, the mutant Kr-NZΔ•86QK was also included. This mutant lacks the C-terminal repression region and contains a specific point mutation in the N-terminal repression region that completely abrogates its ability to inhibit transcription (28). All of the Kr proteins contain the zinc finger region and would therefore bind the Kr sites in the reporter construct. We examined in detail the effects of these repressors on four activators (Fig. 5B): Gal4, an acidic activator; Gal4-Q, a fusion of the Gal4 DNA-binding domain to the second glutamine-rich activation domain of Sp1; Gal4-AH, a fusion of the Gal4 DNA-binding domain to a synthetic acidic amphipathic α-helical peptide that functions as an activation domain; and Gal4-HOX, a fusion of the Gal4 DNA-binding domain to an N-terminal activation region from the mouse *Hox1.3* gene (see Materials and Methods).

We first examined gene expression mediated by the activator Gal4-Q, in the presence of increasing amounts of each repressor construct (Fig. 6A). Consistent with our previous results (29), Kr-NZC efficiently repressed transcription mediated by this activator in a dose-dependent manner (Fig. 6A). Both Kr-ΔZC and Kr-NZΔ also repressed transcription mediated by this activator (Fig. 6A). As expected, the activity of the N-terminal repression region was specifically eliminated by mutation of glutamine 86 to a lysine (Fig. 6A, Kr-NZΔ•86QK). Furthermore, the ability of the Kr derivatives to repress transcription mediated by Gal4-Q was dependent on the presence of DNA-binding sites for Kr (data not shown). In a similar manner, each of the two repression domains of Kr, as well as the full-length Kr protein, repressed transcription mediated by a quite distinct transcriptional activator, Gal4-AH (Fig. 6B).

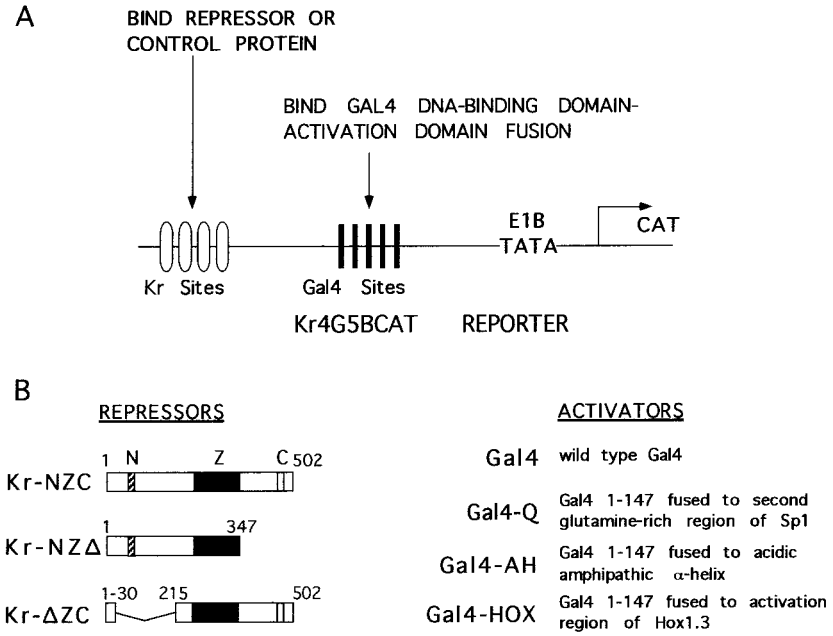


FIG. 5. Scheme of the activator specificity assay. (A) The synthetic promoter used in these assays is depicted. Open ovals indicate the four DNA-binding sites for Kr, and filled rectangles indicate the five DNA-binding sites for Gal4. The promoter contains the TATA box from the adenovirus E1B gene and drives expression of the CAT reporter gene. Gal4 DNA-binding domain–activation domain fusion proteins were expressed from the SV40 early promoter. These activator proteins activated the reporter gene through the Gal4 DNA-binding sites. Activated levels of CAT gene expression were then compared in the presence of either a control protein or Kr derivatives expressed from the simian cytomegalovirus promoter. (B) Kr derivative repressor proteins are depicted, and activator proteins are described. The filled portion of the Kr sequence indicates the DNA-binding domain, marked Z. The hatched region indicates the position of the N-terminal repression region, marked N. The shaded region indicates the position of the C-terminal repression region, marked C.

**Gal4-HOX- or Gal4-mediated transcription is repressed by the N-terminal repression region of Kr only.** In contrast to the results obtained with Gal4-Q- and Gal4-AH-mediated activation, Kr-ΔZC did not repress transcription mediated by Gal4-HOX (Fig. 6C), demonstrating activator specificity of the C-terminal repression region. However, Kr-NZΔ was an efficient repressor of Gal4-HOX-mediated transcription (Fig. 6C), with repression being appropriately relieved by the glutamine-to-lysine mutation at residue 86. The full-length protein, containing both repression regions, repressed Gal4-HOX-mediated transcription, but not as effectively as the N-terminal region alone (Fig. 6C, Kr-NZC); therefore the combination of the two repression regions resulted in an averaged phenotype.

Similarly, although the C-terminal repression region did not repress transcription mediated by the acidic activator Gal4 (Fig. 6D, Kr-ΔZC), the N-terminal repression region of Kr was able to reduce the level of transcriptional activation by Gal4 (Fig. 6D, Kr-NZΔ). Again, the point mutation in the N-terminal domain specifically eliminated this repression activity (Fig. 6D, Kr-NZΔ•86QK). As previously published (29), full-length Kr, like Kr-ΔZC, containing the C-terminal repression region alone, did not repress Gal4-mediated transcription (Fig. 6D, Kr-NZC).

Several aspects of these experiments argue against trivial explanations of the activator specificity of the C-terminal region. First, the abilities of the repression domains to repress the activators were tested over similar ranges of activation activity (see the legend to Fig. 6 for values). Second, by Western blot analysis and gel mobility shift analysis of transfected cell extracts, Gal4-Q, whose stimulatory activity was repressed by Kr-ΔZC, was substantially more highly expressed than the other three activators (data not shown). Thus, neither the expression level (and occupancy on the promoter) nor the strength of the activator could be correlated, even in these limited cases, with the ability of the C-terminal region to inhibit transcription.

Activation mediated by two other activator proteins, a Gal4-VP16 fusion and a fusion of the Gal4 DNA-binding domain to the acidic activation region of EBNA2 (50), was also analyzed. Neither repression region of Kr appeared capable of suppressing transcription mediated by either of these activators (data not shown).

## DISCUSSION

DNA-binding transcriptional repressors exert their effects through a variety of mechanisms. Some mechanisms involve DNA-binding competition and steric hindrance whereby the repressor, by binding to its specific DNA-binding site, blocks the access of basal transcription factors or transcriptional activators to the promoter. Other mechanisms require the repressor to bind to the DNA and perform an active function. Two proposed modes of the latter type of transcriptional repression are termed direct repression and quenching (26). Direct repressors interfere with the formation or activity of the basal transcription complex, whereas quenching repressors interfere with the stimulatory activity of an activator that is also bound to the promoter. The Kr protein is believed to act by one of these active mechanisms, as Kr repression activity does not depend on the position or the sequence of the DNA-binding site through which it is tethered to the DNA (27, 29, 41, 57). Our data described above indicate that Kr contains two small, highly conserved, discrete repression regions, with distinct abilities to repress transcription. These results imply that Kr may have distinct transcriptional targets, perhaps for dis-

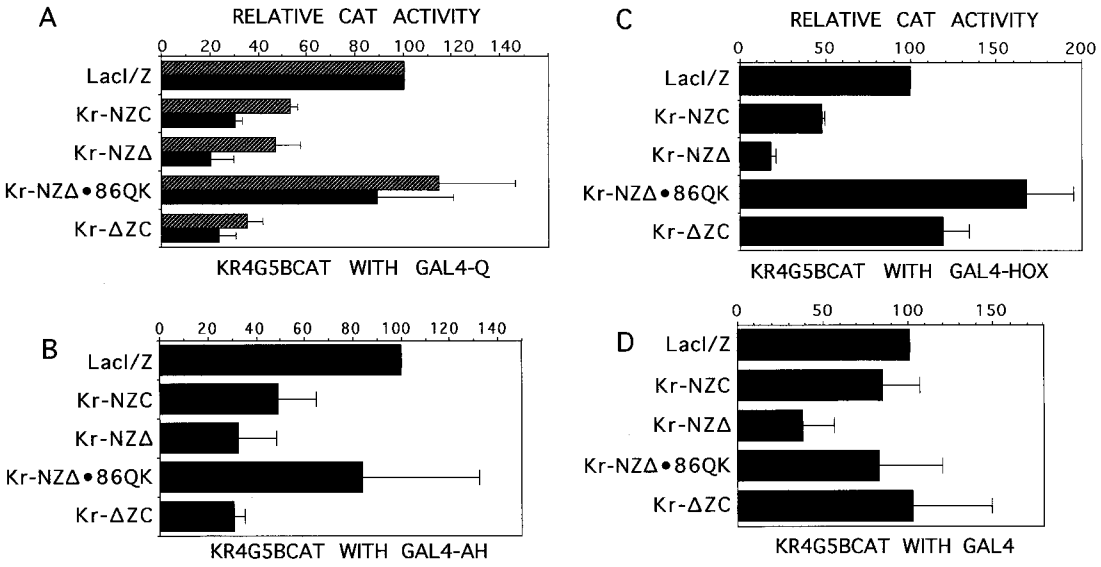


FIG. 6. The N- and C-terminal repression regions in Kr differ in activator specificity. (A) Kr4G5BCAT (2 μg) was cotransfected into CV-1 cells with 5 to 10 ng of Gal4-Q expression vector, 1.0 μg (hatched bars) or 2.0 μg (filled bars) of each of the Kr or Kr derivative expression plasmids, or a control expression plasmid (CMV-LacI/Z) and 0.5 μg of internal control tkGH plasmid, using a LipofectAMINE transfection protocol. CAT activities (determined as percent acetylated chloramphenicol) for each sample were corrected for the level of hGH in the cell media to normalize for transfection efficiency. Relative CAT activity was determined by normalizing corrected CAT activity for each sample to the corrected CAT activity in the presence of the control LacI/Z plasmid. Indicated activities are the averages of three experiments performed in duplicate. Fold activation by Gal4-Q ranged from 50 to 170. (B) Kr4G5BCAT (2 μg) was cotransfected into CV-1 cells with 5 to 10 ng of Gal4-AH expression vector, 1.0 μg of each of the Kr or Kr derivative expression plasmids or a control expression plasmid (CMV-LacI/Z), and 0.5 μg of internal control tkGH plasmid, using a LipofectAMINE transfection protocol. Relative CAT activities were determined as described above. Fold activation by Gal4-AH ranged from 3 to 29. (C) Kr4G5BCAT (2 μg) was cotransfected into CV-1 cells with 5 to 50 ng of Gal4-HOX expression vector, 1.0 μg of each of the Kr or Kr derivative expression plasmids or a control expression plasmid (CMV-LacI/Z), and 0.5 μg of internal control tkGH plasmid, using a LipofectAMINE transfection protocol. CAT activities (determined as percent acetylated chloramphenicol) for each sample were corrected for the level of hGH in the cell medium to normalize for transfection efficiency. Relative CAT activities were determined as described above. Indicated activities are the average of three experiments performed in duplicate. Fold activation by Gal4-HOX ranged from 28 to 250. (D) Kr4G5BCAT (2 μg) was cotransfected into CV-1 cells with 5 to 10 ng of Gal4 expression vector, 1.0 μg of each of the Kr or Kr derivative expression plasmids or a control expression plasmid (CMV-LacI/Z), and 0.5 μg of internal control tkGH plasmid, using a LipofectAMINE transfection protocol. CAT activities were determined as described above. Indicated activities are the average of four experiments, three of which were performed in duplicate. Fold activation by Gal4 ranged from 9 to 141.

tinct biological functions during the course of *Drosophila* development.

**Kr contains two highly conserved, discrete repression regions.** The C-terminal region of Kr was previously found to repress transcription in *Drosophila* Schneider cells when tethered to the promoter via its own or a heterologous DNA-binding domain (40, 41, 57). This C-terminal region also repressed transcriptional activation by hunchback (hb) protein, without a requirement for being tethered to the promoter, possibly by mediating the formation of a complex between the Kr and hb proteins (42, 57). As shown above, the C-terminal region of Kr also confers repression activity to a heterologous DNA-binding protein in mammalian cells. Furthermore, both the N-terminal (27, 28) and C-terminal (see below) small, well-defined regions of the Kr protein characterized in mammalian cells as transferable repression domains must be important for the biological function of Kr during *Drosophila* development, due to their strict conservation during 60 million to 80 million years of evolution (Fig. 1 and 2).

A highly conserved, 42-residue region, from amino acids 433 to 474, exhibited as pronounced repression activity as the entire C-terminal region. Previously, we mapped the N-terminal repression region to an even smaller 31-amino-acid region (28). The small size of these repression domains suggests that each region contacts only a single, discrete target to maximally inhibit gene expression (19). In contrast, transcriptional activation domains often map over large regions and appear to require multiple contacts with the transcriptional machinery in order to fully stimulate initiation of transcription (4).

**Amino acid composition of Kr repression regions.** Although charged residues, and basic residues in particular, have been implicated in the function of other repression domains (7, 33, 38, 55), none of these previously identified charged domains showed significant similarity to the small C-terminal repression domain of Kr. Furthermore, mutation of basic residues in the C-terminal repression region did not diminish its ability to inhibit transcription. Other repression regions have been found to be proline rich (9, 18, 31, 54), and 20% of the residues in the C-terminal repression domain of Kr are prolines, implying a connection. Yet our data also indicated that these residues play no role in repression activity of this region. Surprisingly, analysis of multiple amino acid substitution mutations in the C-terminal repression region indicated that the acidic amino acids enhance the transcriptional inhibition. However, none of the multiple point mutations that were constructed abolished repression activity, indicating that the amino acids most critical for contact between Kr and its molecular target have not yet been delineated. In this regard, the C-terminal transcriptional repression activity of Kr resembles the transcriptional activation activity of VP16. Whereas general acidity may play a role in transcriptional activation by VP16, mutation of a critical phenylalanine residue was required to abolish function (35, 53).

The two repression domains in the Kr protein show no significant homology to each other, and different types of residues appear to be important to the function of each region. Presumably, each type of repression domain represents a unique protein-protein interaction surface necessary for inter-

action with a target within the transcriptional machinery that leads to repression. The wide variety of transcriptional repression domains is consistent with there being many targets within the transcriptional machinery where repressors can interfere: DNA-binding activators, coactivator proteins, or basal transcription factors.

**The two repression regions of Kr differ in activator specificity.** To determine whether the two repression regions in Kr functioned by similar or distinct mechanisms, we tested separately the activator specificities of these repression regions. Although both could inhibit transcriptional activation by a subset of activation domains, other activation domains were inhibited by the N-terminal, but not the C-terminal, region of Kr. Therefore, the target of the N-terminal repression region is likely to be a factor more generally required for transcriptional stimulation than the target of the C-terminal repression region. Stated differently, the limited specificity of the C-terminal repression region suggests that this region quenches only certain activators (see below). Interestingly, not all acidic activators were dealt with in a uniform manner by the repression regions of Kr. For instance, the C-terminal repression region repressed transcription mediated by Gal4-AH but not by Gal4. Therefore, this type of activator specificity assay, in addition to characterizing the biological potential of a repressor, may distinguish among members of a class of activators.

Each of the two repression regions of Kr was capable of individually exerting the same level of inhibition on HSV *tk*-driven gene activity as observed in the presence of both (8b). The redundant, rather than additive or synergistic, repression by the two regions of Kr could imply that the two domains contact similar targets. However, the distinct activator specificities of the two repression regions, discussed above, argue against this possibility. Instead, the lack of synergism in repression levels could be explained by postulating that only one of the regions is fully active in the context of the full-length protein under our assay conditions. Consistent with this hypothesis, the N-terminal region repressed both Gal4-mediated and Gal4-HOX-mediated transcription more potently than did the full-length protein (Fig. 6C and D). In the context of the full-length protein, the N-terminal domain is apparently masked.

The predominance of the C-terminal region in the context of the full-length protein is consistent with this Kr repression activity being the most notable in *Drosophila* Schneider cells (40, 57). However, due to the single amino acid difference in this region found upon comparing the protein sequences of *D. melanogaster* and *D. virilis*, the N-terminal domain must also be functional during *Drosophila* development. The masking of the N-terminal domain in our assay could result from a higher affinity of the C-terminal domain for its target, which might competitively interfere with effective interaction of the N-terminal domain with its target. Alternatively, the N-terminal domain may be unavailable for protein-protein interactions due to the conformation of the full-length protein. A mechanism for unmasking the N-terminal repression region may exist in some cells during *Drosophila* development, permitting potent repression by this region.

**Potential distinct biological roles for the two Kr repression regions.** The distinct activities exhibited by the N-terminal and C-terminal repression regions of Kr may relate to the two types of repression by Kr observed previously in *Drosophila* embryos (17). When Kr binding sites were placed near an enhancer, Kr repressed transcription driven by this adjacent enhancer but did not affect transcription from the same promoter driven by an enhancer more distant from the Kr binding sites. In contrast, when Kr binding sites were placed near the start site of transcription, Kr repressed transcription directed by multiple distant enhancers. Gray and Levine proposed two models to explain these transcriptional phenotypes (17). (i) Kr may possess two separate activities, one interfering with activators when bound nearby within an enhancer and the other directly inhibiting basal transcription, thereby blocking multiple activators, when bound at a promoter. (ii) Alternatively, Kr may recruit a repressive complex that functions only locally, via steric hinderance. Our data fit well with the first model, whereby Kr could directly repress basal transcription via the N-terminal repression region and quench activator activity via the C-terminal repression region. Repressors containing both activator specific activity and activator-independent activity, like Kr with its C-terminal and N-terminal repression regions, could contribute to the combinatorial regulation observed in many *Drosophila* segmentation gene enhancers. With Kr bound in an enhancer, its ability to repress transcription mediated by only certain activators could permit enhancer autonomy, preventing Kr bound at one enhancer from affecting activity from neighboring enhancers. In contrast, Kr bound at a promoter region could mediate complete repression of that gene in regions of the embryo where Kr was present.

The ability of the C-terminal repression region of Kr to inhibit activation by only certain activators may, on the surface, be viewed as conflicting with previous data showing that the C terminus of Kr mediates an interaction with the large subunit of basal transcription factor IIE, TFIIEβ (39). However, an interaction with a basal transcription factor would not necessarily eliminate an activator-specific quenching mode of transcriptional repression. The interaction of Kr with TFIIEβ could prevent only certain activator interactions with the basal transcriptional machinery, most simplistically, if this group of activators targets TFIIE, without directly affecting basal transcription.

## ACKNOWLEDGMENTS

We thank Ron Blackman for the *D. virilis* genomic library, Leslie Pick for the *Hox1.3* cDNA, Welcome Bender for the *D. virilis* strain from which we obtained genomic DNA, and Andrew Lassar for suggesting the use of LipofectAMINE reagent.

This work was supported by Public Health Service grant CA-01272, a March of Dimes Basic Research Grant, and a Biomedical Research Support Grant to U.H., NIH grant K11CA01272 to J.D.L., and NIH predoctoral training grant T32-CA09361 supporting W.H.-R.

## REFERENCES

1. **Blackman, R. K., and M. Meselson.** 1986. Interspecific nucleotide sequence comparisons used to identify regulatory and structural features of the *Drosophila* hsp82 gene. J. Mol. Biol. **188:**499–515.
2. **Brent, R., and M. Ptashne.** 1985. A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. Cell **43:**729–736.
3. **Brown, M., J. Figge, U. Hansen, C. Wright, K.-T. Jeang, G. Khoury, D. M. Livingston, and T. M. Roberts.** 1987. *lac* repressor can regulate expression from a hybrid SV40 early promoter containing a *lac* operator in animal cells. Cell **49:**603–612.
4. **Buratowski, S.** 1995. Mechanisms of gene activation. Science **270:**1773–1774.
5. **Carey, M., Y.-S. Lin, M. R. Green, and M. Ptashne.** 1990. A mechanism for synergistic activation of a mammalian gene by GAL4. Nature **345:**361–364.
6. **Cowell, I. G.** 1994. Repression versus activation in the control of gene transcription. Trends Biochem. Sci. **19:**38–42.
7. **Cowell, I. G., and H. C. Hurst.** 1994. Transcriptional repression by the human bZIP factor E4BP4: definition of a minimal repression domain. Nucleic Acids Res. **22:**59–65.
8. **Emili, A., J. Greenblatt, and C. J. Ingles.** 1994. Species-specific interaction of the glutamine-rich activation domains of Sp1 with the TATA box-binding protein. Mol. Cell. Biol. **14:**1582–1593.

8a. **English, M. A., and J. Licht.** Unpublished data.

8b. **English, M. A., J. C. Reddy, and J. Licht.** Unpublished data.

9. **Estruch, F.** 1991. The yeast putative transcriptional repressor RGM1 is a proline-rich zinc finger protein. Nucleic Acids Res. **19:**4873–4877.
10. **Ferreri, K., G. Gill, and M. Montminy.** 1994. The cAMP-regulated tran-

scription factor CREB interacts with a component of the TFIID complex. Proc. Natl. Acad. Sci. USA **91:**1210–1213.
11. **Fibi, M., B. Zinc, M. Kessel, A. M. Colberg-Poley, S. Labeit, H. Lehrach, and P. Gruss.** 1988. Coding sequence and expression of the homeobox gene 1.3. Development **102:**349–359.
12. **Figge, J., C. Wright, C. J. Collins, T. M. Roberts, and D. M. Livingston.** 1988. Stringent regulation of stably integrated chloramphenicol acetyl transferase genes by E. coli lac repressor in monkey cells. Cell **52:**713–722.
13. **Gill, G., E. Pascal, Z. H. Tseng, and R. Tjian.** 1994. A glutamine-rich hydrophobic patch in transcription factor Sp1 contacts the dTAF$_{II}$110 component of the *Drosophila* TFIID complex and mediates transcriptional activation. Proc. Natl. Acad. Sci. USA **91:**192–196.
14. **Giniger, E., and M. Ptashne.** 1987. Transcription in yeast activated by a putative amphipathic α helix linked to a DNA-binding unit. Nature **330:**670–672.
15. **Goodrich, J. A., T. Hoey, C. J. Thut, A. Admon, and R. Tjian.** 1993. Drosophila TAF$_{II}$40 interacts with both a VP16 activation domain and the basal transcription factor TFIIB. Cell **75:**519–530.
16. **Gorman, C. M., L. F. Moffat, and B. H. Howard.** 1982. Recombinant genomes which express chloramphenicol acetyltransferase. Mol. Cell. Biol. **2:**1044–1051.
17. **Gray, S., and M. Levine.** 1996. Short-range transcriptional repressors mediate both quenching and direct repression within complex loci in *Drosophila*. Genes Dev. **10:**700–710.
18. **Han, K., and J. L. Manley.** 1993b. Transcriptional repression by the *Drosophila even-skipped* protein: definition of a minimal repression domain. Genes Dev. **7:**491–503.
19. **Hanna-Rose, W., and U. Hansen.** 1996. Active repression mechanisms of eukaryotic transcriptional repressors. Trends Genet. **12:**229–234.
20. **Hoey, T., R. O. J. Weinzierl, J.-L. C. G. Gill, B. D. Dynlacht, and R. Tjian.** 1993. Molecular cloning and functional analysis of Drosophila TAF110 reveals properties expected of coactivators. Cell **72:**247–260.
21. **Jäckle, H.** Personal communication.
22. **Johnson, A. D.** 1995. The price of repression. Cell **82:**655–658.
23. **Kakidani, H., and M. Ptashne.** 1988. GAL4 activates gene expression in mammalian cells. Cell **52:**161–167.
24. **Kassis, J. A., S. J. Poole, D. K. Wright, and P. H. O'Farrell.** 1986. Sequence conservation in the protein coding and intron regions of the *engrailed* transcription unit. EMBO J. **5:**3583–3589.
25. **Kingston, R. E.** 1987. Calcium phosphate transfection, p. 911–919. *In* F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.), Current protocols in molecular biology, vol. 1. Greene-Wiley Interscience, New York.
26. **Levine, M., and J. L. Manley.** 1989. Transcriptional repression of eukaryotic promoters. Cell **59:**405–408.
27. **Licht, J. D., M. J. Grossel, J. Figge, and U. M. Hansen.** 1990. *Drosophila Krüppel* protein is a transcriptional repressor. Nature **346:**76–79.
28. **Licht, J. D., W. Hanna-Rose, J. C. Reddy, M. A. English, M. Ro, M. Grossel, R. Shaknovich, and U. Hansen.** 1994. Mapping and mutagenesis of the amino-terminal transcriptional repression domain of the *Drosophila Krüppel* protein. Mol. Cell. Biol. **14:**4057–4066.
29. **Licht, J. D., M. Ro, M. A. English, M. Grossel, and U. Hansen.** 1993. Selective repression of transcriptional activators at a distance by the *Drosophila Krüppel* protein. Proc. Natl. Acad. Sci. USA **90:**11361–11365.
30. **Luckow, B., and G. Schütz.** 1987. CAT constructions with multiple unique restriction sites for the functional analysis of eukaryotic promoters and regulatory elements. Nucleic Acids Res. **15:**5490.
31. **Madden, S. L., D. M. Cook, and F. J. Rauscher III.** 1993. A structure-function analysis of transcriptional repression mediated by the WT1, Wilms' tumor suppressor protein. Oncogene **8:**1713–1720.
32. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual, p. 382–389. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
33. **Margolin, J. F., J. R. Friedman, W. K.-H. Meyer, H. Vissing, H.-J. Thiesen, and F. J. Rauscher III.** 1994. Krüppel-associated boxes are potent transcriptional repression domains. Proc. Natl. Acad. Sci. USA **91:**4509–4513.
34. **Ollo, R., and T. Maniatis.** 1987. *Drosophila Krüppel* gene product produced in a baculovirus expression system is a nuclear phosphoprotein that binds to DNA. Proc. Natl. Acad. Sci. USA **84:**5700–5704.
34a. **Pankratz, M. J., M. Hoch, E. Seifert, and H. Jäckle.** 1989. *Krüppel* requirement for *knirps* enhancement reflects overlapping gap gene activities in the *Drosophila* embryo. Nature **341:**337–339.
35. **Reiger, J. L., F. Shen, and S. J. Triezenberg.** 1993. Pattern of aromatic and hydrophobic amino acids critical for one of two subdomains of the VP16 transcriptional activator. Proc. Natl. Acad. Sci. USA **90:**883–887.
36. **Renkawitz, R.** 1990. Transcriptional repression in eukaryotes. Trends Genet. **6:**192–197.
37. **Rosenberg, U. B., C. Schröder, A. Preiss, A. Kienlin, S. Côté, I. Riede, and H. Jäckle.** 1986. Structural homology of the product of the *Drosophila Krüppel* gene with Xenopus transcription factor IIIA. Nature **319:**326–329.
38. **Saha, A., J. M. Brickman, L. Lehming, and M. Ptashne.** 1993. New eukaryotic transcriptional repressors. Nature **363:**648–652.
39. **Sauer, F., J. D. Fondell, Y. Ohkuma, R. G. Roeder, and H. Jäckle.** 1995. Control of transcription by Krüppel through interactions with TFIIB and TFIIEβ. Nature **375:**162–164.
40. **Sauer, F., and H. Jäckle.** 1991. Concentration-dependent transcriptional activation or repression by Krüppel from a single binding site. Nature **353:** 563–566.
41. **Sauer, F., and H. Jäckle.** 1993. Dimerization and the control of transcription by *Krüppel*. Nature **364:**454–457.
42. **Sauer, F., and H. Jäckle.** 1995. Heterodimeric *Drosophila* gap gene protein complexes acting as transcriptional repressors. EMBO J. **14:**4773–4780.
43. **Seldon, R. F., K. Burke-Howie, M. E. Rowe, H. M. Goodman, and D. D. Moore.** 1986. Human growth hormone reporter gene in regulation studies employing transient gene expression. Mol. Cell. Biol. **6:**3173–3179.
44. **Štanojević, D., T. Hoey, and M. Levine.** 1989. Sequence-specific DNA-binding activities of the gap proteins encoded by *hunchback* and *Krüppel* in *Drosophila*. Nature **341:**331–335.
45. **Stringer, K. F., C. J. Ingles, and J. Greenblatt.** 1990. Direct and selective binding of an acidic transcriptional activation domain to the TATA-box factor TFIID. Nature **345:**783–786.
46. **Tautz, D., C. Tautz, D. A. Webb, and G. A. Dover.** 1987. Evolutionary divergence of promoters and spacers in the rDNA family of four Drosophila species. Implications for molecular coevolution in multigene families. J. Mol. Biol. **195:**525–542.
47. **Tautz, D., M. Trick, and G. A. Dover.** 1986. Cryptic simplicity in DNA is a major source of genetic variation. Nature **322:**652–656.
48. **Thut, C. J., J.-L. Chen, R. Klemm, and R. Tjian.** 1995. p53 transcriptional activation mediated by coactivators TAF$_{II}$40 and TAF$_{II}$60. Science **267:**100–104.
49. **Tjian, R., and T. Maniatis.** 1994. Transcriptional activation: a complex puzzle with few easy pieces. Cell **77:**5–8.
50. **Tong, X., F. Wang, C. J. Thut, and E. Keiff.** 1995. The Epstein-Barr virus nuclear protein 2 acidic domain can interact with TFIIB, TAF40, and RPA70 but not with TATA-binding protein. J. Virol. **69:**585–588.
51. **Treier, M., C. Pfeifle, and D. Tautz.** 1989. Comparison of the gap segmentation gene *hunchback* between *Drosophila melanogaster* and *Drosophila virilis* reveals novel modes of evolutionary change. EMBO J. **8:**1517–1525.
52. **Treisman, J., and C. Desplan.** 1989. The products of the Drosophila gap genes hunchback and Krüppel bind to the hunchback promoters. Nature **341:**335–337.
53. **Triezenberg, S. J., R. C. Kingsbury, and S. L. McKnight.** 1988. Functional dissection of VP16, the *trans-activator* of herpes simplex virus immediate early gene expression. Genes Dev. **2:**718–729.
54. **Wang, Z.-Y., Q.-Q. Qiu, M. Gurrieri, J. Huang, and T. F. Deuel.** 1995. WT1, the Wilms' tumor suppressor gene product, represses transcription through an interactive nuclear protein. Oncogene **10:**1243–1247.
55. **Witzgall, R., E. O'Leary, A. Leaf, D. Önaldi, and J. V. Bonventre.** 1994. The Krüppel-associated box-A (KRAB-A) domain of zinc finger proteins mediates transcriptional repression. Proc. Natl. Acad. Sci. USA **91:**4514–4518.
56. **Zhao, J. J., R. A. Lazzarini, and L. Pick.** 1993. The mouse Hox1.3 gene is functionally equivalent to the Drosophila sex combs reduced gene. Genes Dev. **7:**343–354.
57. **Zuo, P., D. Stanojević, J. Colgan, K. Han, M. Levine, and J. L. Manley.** 1991. Activation and repression of transcription by the gap proteins hunchback and Krüppel in cultured Drosophila cells. Genes Dev. **5:**254–264.